NPS 54-80-06

NAVAL POSTGRADUATE SCHOOL

Monterey, California

Performance vs. Paper-and-Pencil Estimates

of Cognitive Abilities

James K. Arima

ared for:

l Postgraduate School
erey, California 93940

NAVAL POSTGRADUATE SCHOOL
Monterey, California

Rear Admiral J. J. Ekelund Superintendent	Jack R. Borsting Provost

Reproduction of all or part of this report is authorized.

This report was prepared by:

Unclassified

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER NPS 54-80- 06	2. GOVT ACCESSION NO.	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) Performance vs. Paper-and-Pencil Estimates of Cognitive Abilities		5. TYPE OF REPORT & PERIOD COVERED Technical Report - Final
		6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s) James K. Arima		8. CONTRACT OR GRANT NUMBER(s)
9. PERFORMING ORGANIZATION NAME AND ADDRESS Naval Postgraduate School Monterey, CA 93940		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS
11. CONTROLLING OFFICE NAME AND ADDRESS None		12. REPORT DATE May 1980
		13. NUMBER OF PAGES
14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office) None		15. SECURITY CLASS. (of this report) Unclassified
		15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release; distribution unlimited.

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

Intelligence	Selection	Minorities	AFQT
Learning	Performance test	Equal Employment	General intelligence
Testing	Culture-free	ASVAB	Differential psychology

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

This study compared a performance-based, culture-free, learning test with a psychometric test battery, the Armed Services Vocational Aptitude Battery (ASVAB), for assessing individual differences among high-school students. Ethnic background (white-nonwhite) and sex were additional parameters. While negligible differences were found on the learning test between groups differentiated on ethnic background or sex, significant differences favoring white and males were found on the Armed Forces Qualification Test (AFQT)--a

DD FORM 1473, 1 JAN 73 EDITION OF 1 NOV 65 IS OBSOLETE

S/N 0102-014-6601

Unclassified

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered) general aptitude composite formed from the ASVAB and used to derive cutting scores for entry into the Armed Services. The best predictor of the learning test score among the 12 tests of the ASVAB and the AFQT composite was the General Information (GI) test. The AFQT was also a reliable predictor, but only for whites. Analysis of the relationships suggested that the learning test was measuring a portion of the large, general intelligence component known to underlie the ASVAB tests and composites. The learning test appeared to measure this pervasive factor better than the AFQT over the entire range of abilities found in this sample. The results suggest that there may be many capable individuals being eliminated in selection programs using psychometric cutting scores of general cognitive ability or intelligence. Those being adversely impacted in this age group are not only the culturally disadvantaged and persons whose first language is not English, but females in general.

TABLE OF CONTENTS

FIGURES

TABLES

PREFACE AND ACKNOWLEDGEMENTS

This was an independent research project of the author that was not supported financially by sources outside of the Naval Postgraduate School, except for the testing equipment. The equipment was purchased from funds provided the author by the Navy Personnel Research and Development Center, San Diego, California, for research in support of its acquisition and initial assignment programs.

I am again indebted to Paul Sparks of the Man-Machine Systems Design Laboratory for creating the test instrumentation. This project could not have been carried out without the participation of LT Vining A. Sherman, Jr., U.S. Navy, who arranged for the subjects and collected and analyzed the data in conjunction with his thesis effort in the Manpower, Personnel, and Training Analysis curriculum at the Naval Postgraduate School. LT Vining also provided helpful insights and valuable suggestions. Mrs. Montez Bunten did her usual superb job in typing the manuscript.

Portions of this paper were presented at the 21st Annual Conference of the Military Testing Association held in San Diego, California, 15-19 October 1979.

PERFORMANCE vs. PAPER-AND-PENCIL ESTIMATES OF COGNITIVE ABILITIES

James K. Arima

Naval Postgraduate School
Monterey, California 93940

INTRODUCTION

The use of traditional, psychometrically created, paper-and-pencil tests for selection has come under considerable criticism in recent times. One dominant source of this critical appraisal is equal employment opportunity legislation and the court decisions that have followed. The tests have been criticized for their cultural bias, and even when they have been shown to be equally valid for various ethnic or socioeconomic groups in the job context, their continued use has been decried on the basis of the adverse impact that results. Another source of criticism has been politically motivated actions capitalizing on the distrust and dislike of objective tests by a segment of the general public. This activity has resulted in the banning of mass testing for pupil classification in California and the so-called "truth in testing" legislation passed in New York (Smith, 1979). Finally, questioning of the construct validity and ecological relevance of factorially developed tests has come from the lack of intersection between test constructs and findings in the rapidly developing area of cognitive psychology (Carroll & Maxwell, 1979; Sternberg, 1979). This last basis for criticism is particularly important to the psychological profession as it points out the distinction made years ago by Cronbach (1957) of the two disciplines of scientific psychology--the correlation and the experimental approaches.

Taking cognizance of these trends, an earlier effort attempted to create a performance test that was practical to administer, had high construct validity, was culture free, and would provide results that could be broadly generalized (Arima, 1978; Young, 1975). In addition, an important consideration in creating the test was to measure an ability that was not being sufficiently assessed by conventional testing procedures and that would simultaneously provide a new dimension for making selection decisions. Accomplishing this could increase the selection pool and provide opportunities for individuals who might have been eliminated by conventional procedures. The new dimension was learning aptitude, defined as the ability to profit from experience. Broadly defined in this manner, learning ability has been proposed as an important indicator of intelligence and that higher levels of intelligence would be demonstrated by the ability to learn a fixed amount of material in a shorter time or a larger amount of material in a fixed period of time (Estes, 1974). Learning ability, manifested by such measures as grade point average, has been frequently used as a dependent variable in traditional test research, but the format and procedures of paper-and-pencil tests have made it impractical to use learning as an independent or selection variable. On the other hand, simple learning tasks have been extensively used and validated in comparative psychology (Bitterman, 1975). Validation in this context has been the demonstration of reliably different levels of performance in humans by age or in animals by the phylogenetic hierarchy (Jensen, 1979).

The test, itself, was a discrimination-learning task in which pairs of random forms were presented sequentially to subjects. One member of a pair was arbitrarily designated as the correct alternative, which the subject learned to identify on the basis of positive reinforcement whenever a correct choice was made. Six different pairs made up a list, and their presentation, a trial. In all, 10 trials were given with the item pairs appearing in different random orders in each of the trials. The test was administered in a machine-paced and a self-paced mode to Navy recruits undergoing basic training.

Significant amounts of learning took place over the 10 trials, and the correlation between odd and even trials showed a reliability of .838 when corrected for a test of full length using the Spearman-Brown formula. There was a low, but significant correlation ($r = .27$, $N = 137$) between the discrimination-learning test scores and the Armed Forces Qualification Test (AFQT) scores attained by the subjects in their entrance testing. When the total group was split into white and nonwhite subjects, only the correlation for the white subjects ($r = .223$, $N = 104$) reached statistical significance at the .05 level. Thus, it appeared that the performance measure might be giving an assessment of the true capability of the nonwhite subjects which the verbal AFQT score failed to accomplish. Since, however, the correlation was .213 for the 33 nonwhite subjects, its lack of significance might have been due to smaller sample size. There was also a significant difference on the learning test between white and nonwhite subjects using the machine-paced mode, but not in the self-paced mode. However, the interaction term of ethnic grouping and presentation mode had a probability between .10 and .20 in the analysis of variance of learning test scores, so the differential effects of presentation mode for the racial groupings was not fully confirmed.

The present effort was a continuation of the original project that was motivated by several reasons. First, the learning test was reconfigured to make it more portable and simple to administer. It was made into a self-paced mode using a correction procedure so that selection of only the correct alternative automatically advanced the test to the next pair of items. These changes required a tryout and comparison of the results with the previous findings. There was a desire to see if the lack of a difference in performance between whites and nonwhites would hold up in the self-paced mode using the reconfigured test. There was also a severe restriction in range in the earlier study because the subjects had been selected for service using the AFQT score as a screen. An unselected group was desired for whom the scores of the entire selection battery would be available for comparison with the discrimination-learning test score.

METHOD

Test Modifications

The test, as developed for the original study (Arima, 1978), had three stimulus "lists" that were presented to individuals and scored by means of a set of "off the shelf" laboratory equipment. It was basically a machine-paced test, and the subject had to press a button to advance the stimuli in the self-paced mode. The equipment was cumbersome and large and required considerable effort to set up. The objectives of the test modifications were to make it simple and portable and to run automatically in a self-paced mode.

Since there was no great effect for similarity of stimuli within or between the lists in the original study, stimulus list 1 from the original study was selected. This list (Fig. 1) was constructed to have the least amount of similarity between the stimuli in each pair and among the pairs of the list. One member of each pair was randomly designated as the correct choice.

The basic equipment for the reconfigured test was an SR-400 Stimulus-Response (S-R) Programmer made by Behavioral Controls, Inc. (BCI). The SR-400 has four clear-plastic panels that can be used to present visual stimuli and also serve as the response keys. Stimuli are presented by means of a fan-folded continuous strip of paper that can be programmed to control each of the four channels. It is essentially a sophisticated "teaching machine." In this application, only the two central panels were used, and the other two were blacked out and deactivated.

As previously, 10 different versions of the stimulus list were made in which the order of the pairs was different, and each member of a stimulus pair randomly occupied the right or left position an equal number of times over all 10 versions. The 10 lists were connected into one continuous sequence with the restriction that any one pair did not appear back-to-back. The lists were physically created by pasting the appropriate random figures to the designated position (right or left) on a sheet of the continuous, fan-folded paper. Each pair was coded for the correct response by punching the appropriate channel of the control segment of the sheet. This was done for the 60 stimulus pairs that constituted the entire, 10-list sequence.

In operation, the SR-400 was programmed to advance to the next stimulus pair when the correct panel (stimulus) had been pressed. Thus, a correction method was used for the reinforcement--i.e., the subject had to make a correct response before the paper would move. A BSI counter incorporated into the setup through a BCI Four-Choice Auxiliary Control Console cumulated correct and incorrect responses, and a timer mounted on the control console cumulated viewing time. (It did not move during the time the programmer was cycling to a new pair.) A stepping counter was built into the rear of the counter to buzz when six consecutive correct responses were made, but it became unreliable and was not used in test runs. The cycle time between stimulus pairs was 1.4 sec., and the equipment was programmed to stop at the end of the 10-list sequence.

Subjects

Subjects were obtained through three high schools in Monterey County, California, that participated in the high school testing program of the Defense Department. In this program, the Armed Services Vocational Aptitude Battery (ASVAB) is administered as a service without cost to high schools for vocational counseling. The results of the testing go initially to the high school counselor,

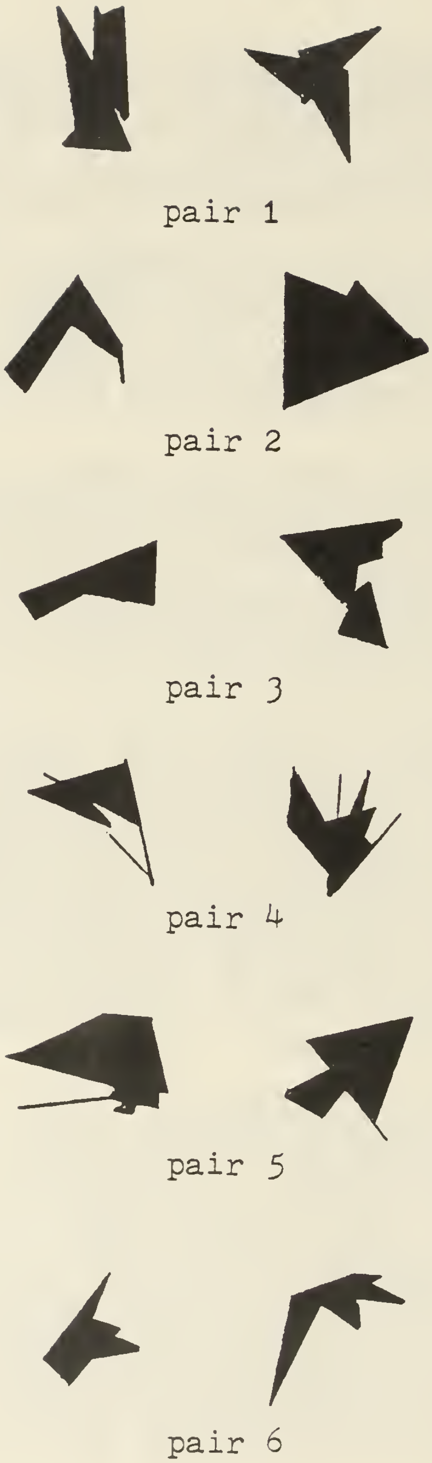

Figure 1. Test Figures

but copies also go to recruiters in the area of the participating schools. Utilizing this source of subjects made it possible to compare learning performance with psychometric test measures in a relatively unselected population, which was one of the purposes of this study. The 65 students with ASVAB scores who were made available for this effort were divided by sex and ethnic grouping as shown in Table 1. The nonwhites were Hispanic (11), black (1), Filipino (2), Oriental (4), Native American (1), and other (3). The subjects came from grades 9 through 12 with the average being 10.7. They ranged in age from 14 through 18 with an average age of 16.2 years.

Table 1

Subjects

Ethnic Group	Male	Female	Total
White	17	26	43
Nonwhite	11	11	22
(Total)	28	37	65

ASVAB

The ASVAB used in the high school testing program was the version identified as Form 5. The tests of the battery, along with their length and reliability, are shown in Table 2. The General Information test includes items of common knowledge that individuals could pick up casually. It was included to provide a measure of the ability of subjects who do not do well in the remainder of the battery, especially those coming from socially deprived environments. Attention to Detail (AD), a perceptual speed test, and Numerical Operations are designed to evaluate potential clerical workers. The Electronic (EI), Shop (SI), and Automotive Information (AI) tests are trade-type tests to identify individuals who already have some capability in these areas or whose familiarity with the material serves as an indication of their interest in this type of work. The other tests are assessments of cognitive skills and stored knowledge. The Armed Forces Qualification Test (AFQT) score is a linear combination of the Word Knowledge (WK), Arithmetic Reasoning (AR), and Space Perception (SP) tests normed on the World War II mobilization population. It has a reliability of .93 (Jensen, et al., 1977). The utilization of the ASVAB in high schools for counseling has been criticized by Cronbach (1979) because it is essentially a selection and placement test as used by the Armed Forces. The Armed Forces Vocational Testing Group has attempted to create composites and provide norms using the relevant population to make it more acceptable for counseling in the high schools while still retaining its primary purpose for the military (U.S. Military Enlisted Processing Command, undated).

Procedure

The test equipment, now quite portable, was set up in the schools where the subjects were available for testing. The instructions were provided to small groups of four or less, but subjects were run in private. The mechanics of the test were explained in the instructions, along with advice that the test was being used for research purposes only and that it was not a timed test but the subject should work quickly without rushing. After the subject's task had been described, they were shown a two-item test not using the figures in the record test to demonstrate how the test would be run and to acquaint the subject with nonsense figures. The subjects were then run individually. Once the first stimulus was presented, the test ran continuously with no apparent break until the 60th frame had been processed.

Table 2

Subtests of the ASVAB Form 5

Name of Test	Number of Items	Subtest Reliabilities*
(GI) General Information	15	.67
(NO) Numerical Operations	50	.88
(AD) Attention to Detail	30	.82
(WK) Word Knowledge	30	.91
(AR) Arithmetic Reasoning	20	.82
(SP) Space Perception	20	.82
(MK) Mathematical Knowledge	20	.88
(EI) Electronic Information	30	.87
(MC) Mechanical Comprehension	20	.81
(GS) General Science	20	.77
(SI) Shop Information	20	.83
(AI) Automotive Information	20	.84

*The data are from Jensen, et al., (1977). The reliabilities were derived using Kuder-Richardson Formula 20 with the exception of Numerical Operations and Attention to Detail, which were obtained by test-retest methods using ASVAB Form 6.

RESULTS

The total exposure time of the stimuli ranged from 35.5 to 161.1 sec. with a mean exposure time of 79.1 sec. Incorporating the 1.4-sec. cycle time between stimuli, the individual administration of the test required an average of 2.7 min. Since all subjects were administered 60 stimulus pairs, those with the shortest exposure times were averaging a little over .5 sec. per frame. Speed on the test could be a characteristic of quick learning or a rapid response set. The latter might be the result of negative motivational factors induced by telling the subjects that the test was being given for strictly research purposes. Questions about the role of rate of responding carried considerable concern, since the scoring of the test was in terms of the number of correct responses per unit of viewing time. This was called the Information Processing Rate (IPR), since each stimulus pair carried one bit of information. The correlation between the number correct and viewing time was -.73, which was significant at the .01 level. This indicated that the individuals who learned more required less time. Accordingly, it was concluded that subjects were motivated to perform well and that quick responding was, as originally hypothesized, an indication of rapid learning.

The means and standard deviation on all subtests of the ASVAB, the AFQT composite, and the IPR are shown in Table 3 by sex, ethnic group, and the entire sample. The IPR was multiplied by 1,000 for convenience in displaying the ratio.

At the .05 significance level, there were no male-female differences in IPR scores for the total sample or the subsamples. There were significant differences between all whites and nonwhites (t = 2.20) and between white and nonwhite females (t = 2.30). The difference of 72.42 in the mean scores of white and nonwhite males did not reach statistical significance. Thus, it appears that there are white-nonwhite differences in IPR performance, and that this difference was due primarily to differences between females of the two groups.

On the AFQT, there was a significant difference in mean scores between males and females at the .05 level for only the white subjects (t = 2.26). No differences were found between all males and females and between nonwhite males and females. There were significant white-nonwhite differences in mean AFQT scores for all categories of subjects. The white-nonwhite difference for all subjects was significant at the .01 level (t = 3.00). The differences between white and nonwhite males (t = 2.44) and between white and nonwhite females (t = 2.10) were significant at the .05 level. To summarize, there are consistent differences between all white and nonwhite groupings on the AFQT dimension. The only sex-related difference occurs between male and female whites.

Because of the differences in the sizes of the subsamples, the t-test was used to assess the differences for each contrast rather than an analysis of variance incorporating all of the variables simultaneously. In the significant differences that were found, the higher mean was always for whites or males.

The correlation of the IPR score with the ASVAB tests and the AFQT composite are shown in Table 4 for the total sample and by sex and ethnic groups. The most noteworthy correlations in Table 4 are those between IPR and General Information (GI) for the total sample and for nonwhites at a significance level of .01 and for females at a significance level of .05. The correlation of IPR with Mechanical Comprehension (MC) followed a similar pattern, except that the correlation was not as high, and for females, the correlation of .31 was significant at only the .06 level. There was also a low but significant correlation of IPR with AFQT for the total sample and females. There is a complete absence of correlation between IPR and the psychometric test variables for whites and males. In the case of the former, General Information (GI) and Automotive Information (AI) are the highest correlations, while General Information and Mechanical

TABLE 3

MEAN TEST SCORES BY RACE AND SEX - ALL SCHOOLS

	WHITE			NONWHITE			TOTAL		
	male	female	total	male	female	total	male	female	total
N	17	26	43	11	11	22	28	37	65
GI	10.29	7.85	8.81	10.09	6.46	8.27	10.21	7.43	8.63
	1.69	1.83	2.13	2.95	2.12	3.12	2.22	1.99	2.50
WK	21.88	17.89	19.46	17.36	13.46	15.41	20.11	16.57	18.09
	5.08	5.60	5.69	7.49	6.79	7.26	6.41	6.23	6.50
MK	14.47	13.15	13.67	10.64	11.73	11.18	12.96	12.73	12.83
	4.19	4.32	4.27	4.43	5.41	4.86	4.62	4.64	4.59
GS	12.06	8.92	10.16	8.91	6.73	7.82	10.82	8.27	9.37
	4.01	3.03	3.74	2.81	3.04	3.07	3.86	3.16	3.68
NO	36.53	36.50	36.51	31.91	36.09	34.00	34.71	36.38	35.66
	7.98	8.05	7.92	9.75	11.40	10.57	8.84	9.00	8.90
AR	13.47	11.96	12.56	10.64	10.00	10.32	12.36	11.38	11.80
	4.24	3.18	3.67	4.23	3.19	3.67	4.39	3.27	3.79
EI	17.24	12.31	14.26	14.64	12.82	13.73	16.21	12.46	14.08
	5.87	4.10	5.39	4.52	2.82	3.80	5.45	3.73	4.88
SI	12.88	8.92	10.49	11.64	6.91	9.27	12.39	8.32	10.08
	3.77	2.56	3.63	3.78	2.17	3.86	3.76	2.59	3.72
AD	13.71	14.73	14.33	14.64	13.09	13.86	14.07	14.24	14.17
	3.87	3.09	3.41	3.33	4.78	4.10	3.63	3.69	3.63
SP	12.41	10.58	11.30	8.46	9.00	8.73	10.86	10.11	10.43
	5.41	3.69	4.48	3.39	4.38	3.83	5.05	3.91	4.42
MC	11.94	8.35	9.77	9.82	5.82	7.82	11.11	7.60	9.11
	3.60	3.05	3.69	2.68	1.17	2.87	3.38	2.86	3.54
AI	9.35	7.15	8.02	8.91	5.82	7.36	9.18	6.76	7.80
	4.89	3.03	3.97	2.91	2.27	3.00	4.16	2.86	3.66
AFQT	47.77	40.42	43.33	36.46	32.46	34.46	43.32	38.05	40.32
	11.36	9.73	10.89	12.45	12.36	12.28	12.87	11.03	12.05
IPR	659.06	648.00	652.37	586.64	450.00	518.32	630.60	589.14	607.00
	248.80	265.64	256.15	176.89	152.52	175.69	222.64	252.75	239.32

Note. Top number is test mean. Bottom number is test standard deviation. The table is from Sherman (1979).

Table 4

Correlations of IPR with ASVAB Tests and AFQT Composite

	ASVAB Subtests**												
Group	GI	WK	MK	GS	NO	AR	EI	SI	AD	SP	MC	AI	AFQT
Combined	.34*	<u>.25</u>	.20	.21	.00	.18	.03	.23	.08	.19	<u>.28</u>	.22	<u>.26</u>
White	.26	.14	.14	.06	-.03	.09	.03	.19	.00	.14	.16	.27	.16
Nonwhite	.50*	.33	.17	<u>.44</u>	-.08	.18	-.05	.26	.26	.08	<u>.48</u>	-.03	.27
Male	.34	.00	.08	.20	.05	.11	.00	.14	.18	.14	.24	.10	.10
Female	<u>.36</u>	<u>.40</u>	.28	.19	-.02	.23	-.01	<u>.32</u>	.03	.22	.31	.31	<u>.37</u>

*Significant at $\underline{p} \leq .01$.

_Significant at $\underline{p} \leq .05$.

(Data are from Sherman, 1979.)

**See Table 3 for full test titles.

Comprehension (MC) are the highest for the males. Thus, the nonwhites and females appear to be the prime contributors to any obtained relationship between the IPR scores and the psychometric test variables.

In order to obtain an indication of the relationship to IPR of all of the variables in the study considered simultaneously, the IPR scores were regressed in a stepwise manner on the study variables using the SPSS program (Nie, et al., 1975). The independent variables included the ASVAB tests, the AFQT composite, two dummy variables for the three high schools, a dummy variable for ethnic group, and a dummy variable for sex. Interactive variables were created by multiplying the General Information and AFQT scores by each of the dummy variables. The stepwise procedure was stopped when the adjusted r^2 did not improve and the significance of the overall *F* ratio for regression failed to improve. The fitted equation is shown in Table 5. The obtained r^2 was .239 (adjusted = .189).

Table 5

Stepwise Regression of IPR on the Study Variables

Variables in Equation*	B	Beta	Std Error B	F	Sig
GI	32.22	.34	11.83	7.42	.01
AFD3	3.48	.33	1.28	7.33	.01
GID1	-12.49	-.21	6.91	3.27	N.S.
EI	-8.67	-.18	6.16	1.98	N.S.
constant 384.96					

*See text for identification of the variables.

With 4, 60 d.f., the obtained *F* ratio of 4.7 for regression was significant at the .005 level. It should be noted, however, that other interpretations of the r^2 in stepwise regression might not consider the obtained r^2 to be statistically significant (Wilkinson, 1979).

The variables in the equation included General Information (GI); an interactive variable, AFQT times D3, the race dummy (1 = white, 0 = nonwhite); GI times a school dummy; and EI (Electronics Information). Only the first two contributed to the equation at a statistically significant level. Thus, for all subjects, GI was the best predictor of IPR and for whites, the AFQT was also a significant predictor. The latter would seem to incorporate the fact that whites scored higher than nonwhites on both the IPR and the AFQT. The latter was the best variable to scale the difference between whites and nonwhites on the IPR.

DISCUSSION

Comparison with Previous Findings

One of the objectives of the study was to compare its findings with the results of the original study using the discrimination learning test (Arima, 1978).

The IPR in the previous study for the self-paced condition was 216.5. The IPR in the present study was 607.0. The possible sources of the difference are too many to make reliable comparisons. However, the two items that stand out are the automation of the present version vs. the manual advance of the earlier test and the correction method (contingent reinforcement) used in the present study. In the present study, the subject had to press the correct alternative to advance the system, whereas the subject in the former study was merely informed by a light when he or she made the correct choice by depressing the appropriate response buttons.

There were significant white-nonwhite differences in the machine-paced condition of the earlier study that apparently disappeared in the self-paced mode. There are still significant white-nonwhite differences, but the primary contribution to this difference comes from the female subjects where there was a 200-point difference favoring the white females. Since there were no sex differences among the white subjects, and there was a 136-point difference between male and female nonwhite subjects (Table 3), it appears that the nonwhite females were a particularly low-performing sample. There were no females in the previous study and no significant difference between white and nonwhite male subjects in the present study. Accordingly, there is some justification for concluding that there are no reliable differences between white and nonwhite male subjects. More data would be required to make a similar statement for the female subjects.

In the earlier study, there was a statistically significant correlation between IPR and the AFQT for the total sample and the white subsample. The correlation was not significant for the nonwhites. In this study, there is still a significant relationship between the IPR and AFQT for the total sample, but the significant subsample correlation now occurs in the female subsample. Nevertheless, in view of the repetition of the significant correlation for the larger (total) sample and the regression equation in which, as formerly, the AFQT plays a significant role for only the white subjects, it is concluded that there is modest, but reliable, relationship between the learning performance and the AFQT score. This relationship is further explored below.

Relationships between Learning Performance and ASVAB Test Scores

The relationship between learning performance and the psychometrically derived ASVAB test scores is of particular interest to this study. There is no doubt that a close relationship exists between the IPR scores and GI (General Information). This is evident in the degree and pattern of correlations seen in Table 4 and in the regression equation in Table 5. As previously stated, GI is a test instigated by the Army to provide a "bottom" to the ASVAB. The Army needed a test to differentiate the potential usefulness of individuals who score low on the basic tests used for screening enlistees. In the present study, the highest correlations between IPR and GI scores occurred for those subsamples scoring lower on the AFQT--nonwhites and females. For subjects scoring higher on the AFQT, it may be that ceiling effects in both variables attenuated the calculated relationship (correlation) between them.

To explore the IPR-GI relationship further, the nature of GI, itself, should be examined. First, Table 3 shows that there is no white-nonwhite difference in the GI scores. This comparison holds up very well when the comparison is made between white and nonwhite males and white and nonwhite females. There appears to be a large difference in GI between males and females, just as there is in the AFQT scores. It is remarkable--considering that the other tests of the ASVAB are longer, more reliable, and generally recognized to be the "heavy-weights" in evaluating individuals--that the 15-item GI test should stand out as the best predictor of learning performance. The relationship of GI to the other ASVAB tests is shown in Table 6. The table reveals that GI is significantly correlated with every other subtest in the battery, and it is also one of three tests that identify the first factor (Verbal) extracted from the test correlations (U.S. Enlisted Processing Command, undated). The factor was identified as the ability to tie words and information together. The foregoing would seem to justify the contention that GI is a measure of a strong general factor that pervades and dominates the ASVAB tests and especially the composites (Cronbach, 1979).

Table 6

CORRELATION BETWEEN GENERAL INFORMATION AND OTHER ASVAB SUBTESTS*

NO	AD	WK	AR	SP	MK	EI	MC	GS	SI	AI
44	27	61	52	34	52	61	57	59	61	57
28	14	52	47	34	43	53	51	49	50	47

*Based on Service standardization sample (upper row) and sample of 2,052 students in the 10th, 11th and 12th grades (bottom row).

As for the IPR score, it has only been identified as a rote learning score. It is not a perceptual or speed test, as evidenced by the zero or near-zero correlations between it and the Attention to Detail and Numerical Operations subtests. It is also not dependent on spatial perception as demonstrated by a relatively low correlation with the SP test in Table 4. It is correlated, for the general sample, with Word Knowledge, Mechanical Comprehension, and the AFQT. On the basis of the differential test results, the IPR score is apparently the result of coding (labeling), organizing, and storing in short-term memory for immediate retrieval discriminating information about the nonsense form, stimulus pairs. Jensen (1979) states that this sort of a task makes moderate demands on the concept he calls _g_, a general measure of mental ability or intelligence.

From the preceding analysis of the characteristics of the GI test and the IPR measure, it is hypothesized that they are both measuring a general capacity for processing and using information and a general characteristic of alertness and responsivity to the environment. One would conjecture that either measure would be related to the latency of the alerting response as measured in recent studies using averaged brain potential responses to a light stimulus. These concepts require experimental verification, of course.

As a performance measure, the learning task could be improved and made more discriminating of individual differences if an individually determined stopping criterion were used. For example, 12 successive, correct responses might be such a criterion. The intent was to investigate this possibility, but the instrumentation proved to be uncooperative. A fixed number of trials, as well as paced presentations, penalizes the rapid learner. The information processing rate should be calculated for the learning period and not attenuated by the time required for reflexive responding once the material has been learned.

Implications for Personnel Selection

If the IPR were scored with an individual stopping criterion in order to increase the variance in performance among individuals, it would seem to be an effective and efficient measure of the general intelligence of a person that is reasonably culture free. While it apparently measures the same area as a general factor that dominates the ASVAB, it provides the opportunity for those with poorer language skills to show their capabilities in the areas of the highly language-dominated tests of the ASVAB. With the advent of computerized testing, this and similar performance tests should be simple and efficient to administer and could provide a greater pool of individuals for selection. Moreover, there has been little validation of the selection instruments with performance in the Armed Services because positive correlations are typically not found. It could be that simple performance tests used as selectors might provide the dimensions to better the validation of selection tests. In the area of truth in testing, the performance tests would have a great advantage since the correct answers could be tailored for each subject at the time of testing if the tasks were designed to permit this option. For example, in the present discrimination learning test, the correct member of each pair could be randomly determined immediately prior to testing.

If the ASVAB composites are so dominated by the general factor to make them essentially useless for counseling as asserted by Cronbach (1979), the same could be said for their use in placement, as employed by the Armed Services. Reliable differences must exist between the composites to make either function possible. Unfortunately, the correlations among the key technical Navy composites range from .88 to .91. Swanson (1978) provides validation data for end-of-course grades or time-to-completion of self-paced courses for 19 schools using the General Technical Composite and 8 schools using the Mechanical composite. In almost all of the cases, the correlations are higher for the Electronics composite. The Electronics composite holds up well as the selector with the highest correlation for the 9 schools using it as a selector. Judging from these limited examples, it would be more efficient just to use the Electronics composite as the selector for all of the schools shown in Swanson's study. This study has served to reinforce the notion of a general factor dominating the ASVAB tests by calling attention to the pervasive relationship of General Information to all of the tests and the fact that the General Information test best predicts scores on a discrimination-learning, performance test.

Finally, attention should be called to the case of the females in this study. They are typical of standardization populations in general for the

ASVAB (Jensen, 1977) in that their AFQT scores are one-half a standard deviation lower than the males, and they do poorly in the trade tests. If the standardization norms are strictly applied, the females are very adversely affected in selection for service or the more desirable technical courses. They maintain equity only in the areas of Attention to Detail and Numerical Operations that are the key elements of the Clerical composite. It should be noted again that the mean IPR scores of the white males and females were identical, indicating that they were comparable in general cognitive ability.

REFERENCES

Arima, J. K. A culture-free performance test of learning aptitude (NPS-54-78-2). JSAS Catalog of Selected Documents in Psychology, 1978, 8, 93. (MS 1775)

Bitterman, M. E. The evolution of intelligence. Scientific American, January 1965, 92-100.

Bitterman, M. E. The comparative analysis of learning. Science, 1975, 188, 699-710.

Carroll, J. B., and S. E. Maxwell. Individual differences in cognitive abilities. Annual Review of Psychology, 1979, 30, 603-640.

Cronbach, L. J. The two disciplines of scientific psychology. American Psychologist, 1957, 12, 671-684.

Cronbach, L. J. The Armed Services Vocational Aptitude Battery--A test battery in transition. Personnel and Guidance Journal, 1979, 232-237.

Estes, W. K. Learning theory and intelligence. American Psychologist, 1974, 29 740-749.

Jensen, A. R. The nature of intelligence and its relation to learning. Journal of Research and Development in Education, 1979, 12, 79-95.

Jensen, H. E., I. H. Massey, and L. D. Valentine, Jr. Armed Services Vocational Aptitude Battery Development (ASVAB Forms 5, 6, and 7) (Technical Research Note 77-3). Fort Sheridan, Illinois: U.S. Army Enlisted Processing Command, 1977.

Kettner, N. Armed Services Vocational Aptitude Battery (ASVAB Form 5): Comparison with GATB and DAT tests (Technical Research Report 77-1). Fort Sheridan, Illinois: U.S. Army Enlisted Processing Command, 1977.

Nie, N. H., C. H. Hull, J. G. Jenkins, K. Steinbrenner & D. H. Bent. Statistical package for the social sciences. (2nd ed.) New York: McGraw-Hill, 1975.

Sherman, V. A., Jr. Culture-free testing for selection of recruits. Unpublished master's thesis, Naval Postgraduate School, 1979.

Smith, R. J. "Truth-in-testing" attracts diverse support. Science, 1979, 205, 1110-1114.

Sternberg, R. J. The nature of mental abilities. American Psychologist, 1979, 34, 214-230.

Swanson, L. Armed Services Vocational Aptitude Battery Forms 6 and 7: validation against school performance--interim report (NPRDC TR 78-24). San Diego, Calif.: Navy Personnel Research and Development Center, 1978.

U.S. Military Enlisted Processing Command. ASVAB counselor's guide. Fort Sheridan, Illinois: Author, undated.

Wilkinson, L. Tests of significance in stepwise regression. Psychological Bulletin, 1979, 86, 168-173.

Young, P. A. A culture-free performance test of general learning ability. Unpublished master's thesis, Naval Postgraduate School, Monterey, California, December 1975.

DISTRIBUTION LIST

DEFENSE TECH. INFORMATION CENTER
CAMERON STATION
ALEXANDRIA, VA
22314 (2)

DEAN OF REASEARCH, CODE 012
NAVAL POSTGRADUATE SCHOOL
MONTEREY, CA
93940 (1)

LIBRARY, CODE 0212
NAVAL POSTGRADUATE SCHOOL
MONTEREY, CA
93940 (2)

LIBRARY, CODE 54
NAVAL POSTGRADUATE SCHOOL
MONTEREY, CA
93940 (1)

DEPARTMENT CHAIRMAN, CODE 54
NAVAL POSTGRADUATE SCHOOL
MONTEREY, CA
93940 (1)

JAMES K. ARIMA, CODE 54AA
NAVAL POSTGRADUATE SCHOOL
MONTEREY, CA
93940 (30)

RONALD A. WEITZMAN (CODE 54WZ)
NAVAL POSTGRADUATE SCHOOL
MONTEREY, CA
93940 (1)

RICHARD S. ELSTER, CODE 54EA
NAVAL POSTGRADUATE SCHOOL
MONTEREY, CA
93940 (1)

JOHN D. SENGER, CODE 54SE
NAVAL POSTGRADUATE SCHOOL
MONTEREY, CA
93940 (1)

SCIENTIFIC DIRECTOR
OFFICE OF NAVAL RESEARCH
SCIENTIFIC LIAISON GROUP
AMERICAN EMBASSY TOKYO
APO SAN FRANCISCO, CA
96503 (1)

DR. GLEN L. BRYAN (CODE 450)
OFFICE OF NAVAL RESEARCH
ARLINGTON, VA
22217 (1)

DR. MARTIN A. TOLCOTT (CODE 455)
OFFICE OF NAVAL RESEARCH
ARLINGTON, VA
22217 (1)

DR. BERT T. KING (CODE 452)
OFFICE OF NAVAL RESEARCH
ARLINGTON, VA
22217 (1)

DR. MARSHALL J. FARR (CODE 458)
OFFICE OF NAVAL RESEARCH
ARLINGTON, VA.
22217 (1)

DR. JAMES LESTER
ONR BRANCH OFFICE
495 SUMMER STREET
BOSTON, MA
02210 (1)

DR. EUGENE GLOYE
ONR BRANCH OFFICE
1030 EAST GREEN ST.
PASADENA, CA.
91101 (1)

DR. CHARLES E. DAVIS
ONR BRANCH OFFICE
536 S. CLARK ST.
CHICAGO, IL.
60605 (1)

PROGRAM COORDINATOR
BUREAU OF MEDICINE & SURGERY
(CODE 71G)
DEPT. OF THE NAVY
WASHINGTON, DC
20390 (1)

BEHAVIORAL SCIENCES DEPARTMENT
NAVAL MEDICAL RESEARCH INSTITUTE
NATIONAL NAVAL MEDICAL CENTER
BETHESDA, MD
20014 (1)

DR. JOSEPH L. YOUNG, DIRECTOR
MEMORY & COGNITIVE PROC.
NATIONAL SCIENCE FOUNDATION
WASHINGTON, DC
20550 (1)

DR. THOMAS G. STICHT
BASIC SKILLS PROGRAM
NATIONAL INSTITUTE OF EDUCATION
1200 19TH ST. NW
WASHINGTON, DC
20208 (1)

DIRECTOR, RESEARCH & DATA
OSD/MRA&L
RM. 3B919
THE PENTAGON
WASHINGTON, DC
20301 (1)

CHIEF, PSYCHOLOGICAL RESEARCH
HQ. US COAST GUARD
WASHINGTON, DC
20590 (1)

MR. ROBERT SMITH
OP-987E
DEPT. OF THE NAVY
WASHINGTON, DC
20350 (1)

MR. MAURICE CALLAHAN
NODAC (CODE 2)
DEPARTMENT OF THE NAVY
BLDG. 2, WASHINGTON NAVY YARD
WASHINGTON, DC
20374 (1)

SCIENTIFIC ADVISOR (OP-01T)
DCNO(MPT)
DEPT. OF THE NAVY
WASHINGTON, DC
20370 (1)

DR. A. L. SLAFKOSKY
SCIENTIFIC ADVISOR (CODE RD-1)
HQ. US MARINE CORPS
WASHINGTON, DC
20380 (1)

MAJ DAN KUHN
HQMC (MPI-20)
WASHINGTON, DC
20380 (1)

DIRECTOR
OFFICE OF MANPOWER UTILIZATION
HEADQUARTERS
MARINE CORPS
MCB, QUANTICO, VA
22134 (1)

DR. J. DEXTER FLETCHER
CYBERNETICS TECHNOLOGY OFFICE
DARPA
1500 WILSON BLVD.
ARLINGTON, VA
22209 (1)

CDR PAUL R. CHATELIER
OUSDRE
3D129 PENTAGON
WASHINGTON, DC
20301 (1)

SYSTEMS ANALYSIS DIVISION
DEPT. OF THE NAVY OP-964D)
WASHINGTON, DC
20350 (1)

DR. WILLIAM L. MALOY
CODE 01A
NAVAL TRAINING COMMAND
PENSACOLA, FL
32508 (1)

CHIEF, HUMAN FACTORS LAB.
NAVAL TRAINING EQUIPMENT CENTER
(CODE N-215)
ORLANDO, FL
32813 (1)

DR. NORMAN J. KERR
CHIEF OF NAVAL TECH TRAINING
NAVAL AIR STATION MEMPHIS (CODE 75)
MILLINGTON, TN
38054 (1)

DR. ALFRED F. SMODE, DIRECTOR
TRAINING ANALYSIS AND
EVALUATION GROUP
DEPT. OF THE NAVY
ORLANDO, FL
32813 (1)

DR. F. WORTH SCANLAND
CNET (N-5)
NAVAL AIR STATION
PENSACOLA, FL
32508 (1)

DR. MARTIN F. WISKOFF (310)
NAVY PERSONNEL R&D CENTER
SAN DIEGO, CA
92152 (1)

LIBRARY
NAVY PERSONNEL R&D CENTER
SAN DIEGO, CA
92152 (1)

DR. WILLIAM MONTAGUE (304)
NAVY PERSONNEL R&D CENTER
SAN DIEGO, CA
92152 (

DR. JAMES MCGRATH (306)
NAVY PERSONNEL R&D CENTER
SAN DIEGO, CA
92152 (1

DR. PAT FREDERICO (304)
NAVY PERSONNEL R&D CENTER
SAN DIEGO, CA
92152 (

DR JOHN FORD (304)
NAVY PERSONNEL R&D CENTER
SAN DIEGO, CA
92152 (

DR. ED AIKEN (349)
NAVY PERSONNEL R&D CENTER
SAN DIEGO, CA
92152 (

CDR CHARLES J. THEISEN, JR
HUMAN FACTORS ENGINEERING DIV
NAVAL AIR DEVELOP. CENTER
WARMINSTER, PA
18974 (

CDR ROBERT S. KENNEDY
NAVAL AEROSPACE MEDICAL
RESEARCH LAB.
BOX 29407
NEW ORLEANS, LA
70189 (

COMMANDING OFFICER
MEDICAL RESEARCH LABORATORY
U.S. NAVAL SUB BASE, NEW LONDON
GROTON, CT
06340 (

TECHNICAL DIRECTOR
NAVAL HEALTH RESEARCH CENTER
FCB 85122
SAN DIEGO, CA
92138 (1)

DR. ED JOHNSON
US ARMY RESEARCH INSTITUTE
5001 EISENHOWER BLVD.
ALEXANDRIA, VA
22333 (1)

DR. E. RALPH DUSEK
US ARMY RESEARCH INSTITUTE
5001 EISENHOWER AVE.
ALEXANDRIA, VA
22333 (1)

DR. MILTON S. KATZ
US ARMY RESEARCH INSTITUTE
5001 EISENHOWER AVE.
ALEXANDRIA, VA
22333 (1)

DR. RALPH CANTER
US ARMY RESEARCH INSTITUTE
US ARMY ADMINCEN
FT BENJAMIN HARRISON, IN
46216 (1)

DIRECTOR
US ARMY HUMAN
ENGINEERING LABORATORIES
ABERDEEN PROVING GROUND, MD
21005 (1)

TECHNICAL DIRECTOR
AFHRL/TT
LOWRY AFB, CO
80230 (1)

NET LIAISON OFFICER
AFHRL/FT
WILLIAMS AFB, AZ
85224 (1)

TECHNICAL DIRECTOR
AFHRL/FT
WILLIAMS AFB, AZ
85224 (1)

DR. ALFRED R. FREGLY
AFOSR/NL, BLDG 410
BOLLING AFB, DC
20332 (1)

DR. G. A. ECKSTRAND
AFHRL/AS
WRIGHT-PATTERSON AFB, OH
45433 (1)

DR. EARL A. ALLUISI
TECHNICAL DIRECTOR
AFHRL
BROOKS AFB, TX
78235 (1)

CHAIRMAN, BEHAV SCI DEPT
NAVAL COMMAND-MANAGEMENT DIVISION
U.S. NAVAL ACADEMY
LUCE HALL
ANNAPOLIS, MD
21402 (1)

DEPT OF BEHAV SCI AND LEADERSHIP
U.S. AIR FORCE ACADEMY
COLORADO SPRINGS, CO
80840 (1)

LIBRARY
U.S. NAVAL ACADEMY
ANNAPOLIS, MD
21402 (1)

LIBRARY
NAVY WAR COLLEGE
PROVIDENCE, RI
02840 (1)

LIBRARY
NATIONAL DEFENSE UNIVERSITY
FT. MCNAIR
WASHINGTON, DC
20319 (1)

SCHOOL OF LOGISTICS AND MANAGEMENT
AFIT
WRIGHT PATTERSON AFB, OH
45433 (1)

DIRECTOR, PERSONNEL SYSTEMS MGT
U.S. ARMY WAR COLLEGE
CARLISLE BARRACKS, PA
17013 (1)

DEPARTMENT OF DEFENSE MGT STUDIES
INDUST COLLEGE OF THE ARMED FORCES
WASHINGTON, DC
20319 (1)

DEPT OF BEHAV SCI AND LEADERSHIP
UNITED STATES MILITARY ACADEMY
WEST POINT, NY
10996 (1)

LEADERSHIP DEPARTMENT
U.S. COAST GUARD ACADEMY
NEW LONDON, CT
06320 (1)

AIR UNIVERSITY LIBRARY
LSE - 8110
MAXWELL AFB, AL
36112 (1)

OFFICER IN CHARGE
HRMD
DEFENSE RACE RELATIONS INSTITUTE
PATRICK AFB, FL
32927 (1)

LEADERSHIP INSTRUCTION DEPARTMENT
MARINE CORPS EDUCATION CENTER
MARINE CORPS ED & DEV CMD
QUANTICO, VA
22134 (1

COMMANDER, CDTC
P.O. BOX 444
FORT ORD, CA
93941 (1

1 PSYCHOLOGICAL RESEARCH UNIT
DEPT. OF DEFENS. (ARMY OFFICE)
CAMPBELL PARK OFFICES
CANBERRA ACT 26[illegible]
AUSTRALIA (1

DIRECTOR
ARMY PERSONNEL RESEARCH
ESTABLISHMENT
FARNBOROUGH, HANTS,
UNITED KINGDOM (1

LIBRARY
CENTER FOR ADVANCED STUDY
IN BEHAVIORAL SCIENCES
202 JUNIPERO SERRA BLVD.
STANFORD, CA
94305 (1

EDUCATIONAL TESTING SERVICE
PRINCETON, NJ
08541 (1

DR. MALCOLM D. ARNOULT
DEPT. OF PSYCHOLOGY
TEXAS CHRISTIAN UNIVERSITY
FT WORTH, TX
76129 (1

DR. NICHOLAS A. BOND
DEPT. OF PSYCHOLOGY
SACRAMENTO STATE COLLEGE
600 JAY ST
SACRAMENTO, CA
95819 (1

DR. LYLE BOURNE
DEPT. OF PSYCHOLOGY
UNIVERSITY OF COLORADO
BOULDER, CO
80302 (1)

DR. JOHN B. CARROLL
PSYCHOMETRIC LAB
UNIV. OF NO. CAROLINA
DAVIE HALL 013A
CHAPEL HILL, NC
27514 (1)

DR. RAYMOND E. CHRISTAL
239 KILLARNEY
SAN ANTONIO, TX
78223 (1)

DR. LEE J. CRONBACH
16 LABURNUM RD.
ATHERTON, CA
94205 (1)

DR. WILLIAM K. ESTES
ROCKEFELLER UNIVERSITY
NEW YORK, NY
10021 (1)

DR. EDWIN A. FLEISHMAN
ADVANCED RESEARCH RESOURCES
ORGANIZATION
8555 SIXTEENTH ST.
SILVER SPRING, MD
20910 (1)

DR. ROBERT GLASER
LRDC
UNIV. OF PITTSBURGH
3939 O'HARA STREET
PITTSBURGH, PA
15213 (1)

DR. DONALD ROSS GREEN
CTB/MCGRAW-HILL
MONTEREY, CA
93940 (1)

DR. WINFRED F. HILL
DEPT. OF PSYCHOLOGY
NORTHWESTERN UNIVERSITY
EVANSTON, IL
60201 (1)

DR. LLOYD G. HUMPHREYS
PSYCHOLOGY BUILDING
UNIVERSITY OF ILLINOIS
CHAMPAIGN, IL
61820 (1)

DR. EARL HUNT
DEPARTMENT OF PSYCHOLOGY
UNIV. OF WASHINGTON
SEATTLE, WA
98195 (1)

DR. ARTHUR R. JENSEN
DEPT. OF EDUCATION
UNIVERSITY OF CALIFORNIA
BERKELEY, CA
94704 (1)

DR. ROBERT LAMBERT
EDUCATIONAL TESTING SERVICE
1947 CENTER STREET
BERKELEY, CA
94708 (1)

DR. HOWARD H. MCFANN
MCFANN/GRAY ASSOCIATES
2100 GARDEN RD. SUITE J
MONTEREY, CA
93940 (1)

DR. JESSE ORLANSKY
INSTITUTE FOR DEFENSE ANALYSIS
400 ARMY NAVY DRIVE
ARLINGTON, VA
22202 (1)

DR. ROBERT SEIDEL
INSTRUCTIONAL TECHNOLOGY GRP.
HUMRRO
300 N. WASHINGTON ST.
ARLINGTON, VA
22314 (1)

DR. ARTHUR I. SIEGEL
APPLIED PSYCHOLOGICAL SERVICES
404 E. LANCASTER AVE.
WAYNE, PA
19087 (1)

DR. RICHARD SNOW
SCHOOL OF EDUCATION
STANFORD UNIVERSITY
STANFORD, CA
94305 (1)

DR. ROBERT STERNBERG
DEPT. OF PSYCHOLOGY
YALE UNIVERSITY
BOX 11A, YALE STATION
NEW HAVEN, CT
06520 (1)

DR. PATRICK SUPPES
INSTITUTE FOR MATH. STUDIES IN
THE SOCIAL SCIENCES
STANFORD UNIVERSITY
STANFORD, CA
94305 (1)

DR. KIKUMI TATSUOKA
COMPUTER BASED ED. RSCH. LAB.
252 ENGINEERING RSCH. LAB.
UNIV. OF ILLINOIS
URBANA, IL
61801 (1)

DR. JOHN TAYLOR
HUMRRO/WESTERN DIV
27857 BERWICK DR.
CARMEL, CA
93923